MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2023, 19(5), em2265
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/13159

# Prospective secondary school teachers' knowledge of sampling distribution properties

Nuria Begué [1], Carmen Batanero [2], María Magdalena Gea [3*], Silvia María Valenzuela-Ruiz [3]

[1] Department of Mathematics, University of Zaragoza, Zaragoza, SPAIN
[2] Research Group FQM126, University of Granada, Granada, SPAIN
[3] Department of Mathematics Education, University of Granada, Granada, SPAIN

Received 05 December 2022 ▪ Accepted 19 March 2023

**Abstract**

The aim of the work was to assess prospective Spanish secondary school mathematics teachers' knowledge of sampling representativeness and variability. To achieve this goal, a questionnaire with four items taken from our previous research with secondary education and high school students was proposed to 66 prospective teachers. In each item, participants were asked to provide four values of a binomial distribution, varying the parameters of the distribution in each item. The analysis of the mean and range of the sample of four values provided by the participants, and its comparison with the theoretically normative and acceptable values of the sampling distribution suggests a good knowledge of the sample representativeness, except for one item and an overestimation of the variability for large values of the number of trials in the binomial distribution. The study of the participants' arguments serves to identify some reasoning biases related to sampling distribution. The results are better than those obtained with students in our previous studies.

**Keywords:** knowledge, prospective teachers, representativeness and variability, sampling distribution

## INTRODUCTION

Knowledge about sampling is essential in research, decision-making and other areas of human activity. Burrill and Biehler (2011) and Heitele (1975) considered it a basic stochastic idea because we obtain much scientific knowledge from sampling. Besides, sampling is a link between statistics and probability. Eichler and Vogel (2014) indicate that sampling is receiving a great deal of attention in mathematics education research due to its role in simulation, which is a valuable didactic resource for understanding probability and statistical inference (Koparan, 2022).

In Spain, primary education consists of six different grades (one to six, children from six to 11 year-olds), secondary education includes four different grades (one to four, 12-15 year-olds), while high school consists of two different grades (one and two, 17-18 year-olds).The teaching of sampling begins in Spain in primary education with the identification of a sample as a data set of a larger set, as well as the reflection on the population to which it is possible to apply the conclusions of simple statistical investigations (MEFP, 2022a). Sampling is formally studied in secondary education with the frequentist approach to probability and in work with simulation (both based on sampling).

At this educational stage, its study consists of the concepts of population and sample, sampling representativeness and variability, and elementary sampling techniques (MEFP, 2022b).

This topic expands in the second year of social sciences high school (students aged 17-18), where the sampling distribution of the mean and proportion and confidence intervals for these parameters are introduced (MEFP, 2022b). Moreover, the analysis of the university entrance exams in mathematics applied to social sciences showed that, in the past 12 years, a task on sampling distribution, confidence intervals or statistical tests (based on sampling distribution) was always included (López-Martín et al., 2016).


✉ nbegue@unizar.es ✉ batanero@ugr.es ✉ mmgea@ugr.es **(*Correspondence)** ✉ svalenzuela@ugr.es

**Contribution to the literature**

- This study adds to the literature new results related to prospective teachers' knowledge of sampling variability and representativeness. Using a questionnaire previously given to student we obtain better results, although reasoning biases are still found in the prospective teachers.
- The tasks and analysis method used are a novelty in research related to teachers' understanding of sampling distributions.
- A specific contribution is the comparison of prospective teachers' and students' responses and arguments in providing samples from binomial distributions.

Understanding the sampling distribution involves prior knowledge of many concepts, which may explain the students' difficulties with this topic (Begué et al., 2019; Kula & Kocer, 2020; Makar & Rubin, 2018). Virtually all previous research was conducted with students, including our work with secondary school students (Begué et al., 2018) and high school students (Batanero et al., 2020). Similar reasoning biases were found in these studies, although the proportion of students who showed them decreased as the school year progressed. Consequently, we wonder if prospective teachers have overcome these difficulties due to their mathematical preparation.

To answer this question, the aim of this paper is to assess the mathematical knowledge about sampling representativeness and variability in a sample of prospective Spanish secondary and high school mathematics teachers. Since most previous research on sampling was carried out with students, our results provide new knowledge concerning prospective teachers' understanding of sampling.

## THEORETICAL BACKGROUND

We base our work on three kinds of foundations:

1. the sampling distribution properties,
2. the idea of mathematical content knowledge for teaching, and
3. previous research.

### Sampling Distribution

The main difficulty of sampling is the need to use, differentiate, and relate three types of probability distributions (Harradine et al., 2011):

1. **The distribution of the population variable:** In our study, the binomial distribution, which represents the number of times that a given event (success) of probability $p$ occurs over $n$ independent repetitions of a random experiment and which is denoted by $B$ ($n$, $p$), where $p$ and $n$ are the parameters of this distribution.
2. **The distribution of the same variable in a random sample of independent elements of that population:** Thus, in our study, the proportion $\hat{p}$ (cases in the sample that meet the given condition) is a statistic that allows estimating the population parameter $p$.
3. **The sampling distribution:** Considering repeated samples of the same size, the statistic obtained in each sample is a random variable that varies from a sample to another. Its distribution is called the sampling distribution. In our study, the proportion $\hat{p}$ follows a normal distribution whose mean is the proportion in the population ($p$). This property is known as representativeness and means that the number of expected successes in each sample is $np$. Moreover, the standard deviation of the sampling proportion is $\frac{\sqrt{p(1-p)}}{\sqrt{n}}$, where $n$ is the sample size. If the sample size increases, the sampling variability decreases.

In this study, we want to assess the prospective teachers' understanding of sampling representativeness and variability, and the relation of sampling variability with the sample size. The results will complement previous research on prospective teachers' understanding of other inferential topics. For example, Valenzuela-Ruiz et al. (2023) studied prospective teachers' competence to solve formal mathematical problems on the sampling distribution and to predict possible difficulties of the students when solving the problems, and Lugo-Armenta and Pino-Fan (2021) analyzed prospective teachers' knowledge learning of the Chi-squared statistics in a virtual environment.

### Mathematical Knowledge for Teaching

There is extensive research on mathematics teacher education, for example, Alshehri and Youssef (2022), Llinares (2018, 2021), Ping et al. (2018), and Scheiner et al. (2021).

We base this paper on the mathematical knowledge for teaching (MKT) model, which the authors divide into the following components: common content knowledge (CCK), specialized content knowledge (SCK), knowledge of content and teaching (KCT), and knowledge of content and students (KCS) (Ball et al., 2008; Phelps et al., 2020). Hill et al. (2008) further proposed horizon content knowledge (HCK) and knowledge of content and curriculum (KCC). CCK is the knowledge brought into play by an educated person to solve mathematical problems for which a person with basic knowledge is qualified. SCK describes the teacher's special competence that enables him to plan and develop teaching sequences. HCK refers to the more advanced aspects of the content, which provide insights for the teacher, for example, knowledge of the history or detection of possible errors related to the mathematical ideas underlying the topic.

This paper aims to analyze the prospective high and secondary school teachers' mathematical knowledge of sampling. Specifically, we focus on sampling representativeness and variability, which is part of CCK because it appears in the secondary school curriculum guidelines (MEFP, 2022b) and is taught to students.

## Background

Research on the understanding of sampling has been made almost exclusively with students and has focused on the two main ideas of sampling representativeness and variability, which must be understood and related to each other (Harradine et al., 2011).

The results of Kahneman and Tversky's (1972) research show that most people expect a sample to be very similar to the population from which it comes, so for their decisions under uncertainty, they consider only representativeness without considering variability (see also Erbas & Ocal, 2022).

The authors called this type of reasoning representativeness heuristic, which gives rise to two biases when predicting results in a series of experiments: positive recency and negative recency. In positive recency, subjects extend the tendency shown in the data series obtained (for example, if a family has three male children, they expect the next child will also be male). In negative recency, people try to compensate the future results with those previously obtained to reach an average (following the example, the family expects the next child to be female).

On the other hand, many students have difficulty to conceive the sampling process, where each sample is a particular case of many samples of the same size, whose frequency can be estimated by calculating probabilities (Saldanha & Thompson, 2002). Consequently, students generally interpret probability questions in a non-probabilistic way, trying to give a prediction of the outcome that will occur instead of calculating the probability of the different possible results. This type of reasoning is known as the outcome approach (Konold, 1989).

Shaughnessy et al. (2004) attempted to assess levels in students' understanding of sampling, describing the following:

1. *Idiosyncratic level:* The students only conceive the samples based on non-mathematical aspects, for example, their preferences.
2. *Additive level:* They view the different samples of the population as disjoint subsets, and the relative frequency of each sample is not considered to predict a sampling distribution.
3. *Proportional level:* The student conceives the sample as a part of the population and relates the sample proportions or means to those in the population; however, they do not consider the variability of the sampling distribution.
4. *Distributional level*: The student considers averages or proportion and sampling variability to predict the sampling distribution.

We also rely on Green's (1983) study on the probabilistic reasoning of 11-14-years-old children. The author proposed a problem that reported the result of throwing 100 thumbtacks into the air, having obtained in a previous experience 68 with the point upwards. He asked the children to predict the number of thumbtacks with the point upwards when repeating the experiment.

His results showed that 64% of the children in the sample indicated that the two events (point or head up) were equiprobable. The author interpreted this response as the equiprobability bias (Gauvrit & Morsanyi, 2014; Lecoutre, 1992), which consists of assuming the events of any random experiment to be equiprobable. Gómez-Torres et al. (2016) adapted the same task to study the understanding of sampling by 157 prospective Spanish primary school teachers. The adaptation consisted of asking for four predictions of the experiment outcome, using the mean value of the estimates to assess understanding of sampling representativeness and the range to evaluate the comprehension of sampling variability. When comparing the means with the expected binomial distribution, the authors identified that 33.8% of the subjects presented the equiprobability bias (Gauvrit & Morsanyi, 2014; Lecoutre, 1992), providing samples whose mean value was around 50%.

Begué et al. (2018) used the same task with 157 Spanish secondary education grade 2 (13-years-olds) and 145 grade 4 (15-years-olds) students. The authors proposed four items to the students, the first of which was taken from Gómez-Torres et al. (2016), varying the parameters of the binomial distribution in the other three. They concluded that the students had a good intuition of the expected value, although about 30% showed the equiprobability bias when the proportion was clearly different from 0.5. They also observed a better understanding of sampling variability in small samples, although there was scarce intuitive understanding of the law of large numbers. Batanero et al. (2020) replicated this research with 234 high school grade 2 students (17-18 years old). The authors performed a different analysis of the results, comparing the students' predictions with the theoretical sampling distribution (instead of using the initial binomial distribution). When comparing with Begué et al. (2018),

**Table 1.** Characteristics of the questionnaire items

| | Item 1 | Item 2 | Item 3 | Item 4 |
|---|---|---|---|---|
| Sample size | 100 | 100 | 10 | 10 |
| Probability of the event of interest | 0.68 | 0.50 | 0.50 | 0.70 |
| Expected number of successes | 68 | 50 | 5 | 7 |
| Normative value of the mean of four predictions (interval containing 68% of sample means) | [65.7-70.3] | [47.5-52.5] | [4.2-5.8] | [6.3-7.7] |
| Acceptable value of the mean of four predictions (interval containing 95% of sample means) | [63.3-72.7] | [45-55] | [3.4-6.6] | [5.6-8.4] |
| Normative value of the range of four predictions (interval containing 68% of the ranges) | [5-14] | [6-15] | [2-5] | [2-4] |
| Acceptable value of the range of four predictions (interval containing 95% of the ranges) | [3-18] | [3-20] | [1-6] | [1-6] |

they found better results in all tasks, although there was a high percentage of students with sampling reasoning biases. They also analyzed the arguments given by a part of the students to justify the samples produced.

The current work continues our previous research by replicating Begué et al. (2018) and Batanero et al. (2020) studies with a sample of prospective secondary school mathematics teachers. We investigate whether the improvement observed in the understanding of sampling representativeness and variability according to the school year continues for prospective teachers. We also go deeper into the reasoning biases observed in these studies and analyze whether they are present in the prospective teachers, who unconsciously could transmit them to their students.

## METHODOLOGY

The participating sample consisted of 66 students in a master's degree program, which is mandatory in Spain to become a mathematics teacher in secondary education and high school. In Spain, those who want to become secondary and high school mathematics teachers must pass a competitive competition with two requirements. The first is to have completed a university degree in mathematics, science, or engineering (four years of university study after finishing high school). The second to complete a specific master's degree (university master in secondary and high school teacher education) in the specialty of mathematics. The master's program is oriented to provide didactic competencies to teach mathematics to secondary and high school students. The topics in this master are teaching and learning mathematics, innovation and introduction to research in mathematics education, psychology and sociology, and teaching practices in a high school. Moreover, the prospective teachers must write and defend a master's thesis on a mathematics education topic.

We used a controlled-selection directed sampling (Wu & Thompson, 2020), which included all the participants in that master program in the 2019-2020 academic year at University of Granada. Half the participants had obtained a university degree in mathematics, with four years of studies in pure and applied mathematics (algebra, calculus, statistics, and probability). The remaining had studied other science or engineering degrees with different topics, including but not limited to mathematics. All of them had at least one complete course in statistics in their careers; however, mathematicians usually had two to three courses in statistics.

We proposed these subjects a questionnaire consisting of four items presenting situations that can be described mathematically by a binomial distribution (**Appendix**), with the characteristics displayed in **Table 1**. There were two variables in the questionnaire. The first variable was the sample size: n=100 in the first two items (large samples) and n=10 in the other two (small samples). The second variable was the proportion value: in items 1 and 4, the probability of success should be estimated by the frequentist approach to probability while the events are equiprobable in the other items.

In **Table 1**, we also present the intervals containing 68% of the means and ranges of the four values requested in each item and the acceptable responses corresponding to 95% of the means and ranges. These intervals serve to characterize the responses as normative or optimal and acceptable. We obtained these intervals using the sampling distribution of the sample proportion of four values for each binomial distribution used and simulating the sampling distribution of the ranges (5,000 simulations).

Once the participants' responses were collected, we computed the mean value and the range of the four values provided in each item. We classified these values according to the intervals presented in **Table 1** and other intervals that indicate the presence of biases, which are discussed in the results section. In addition, we performed a qualitative content analysis (Drisko & Maschi, 2016) of the arguments provided by the prospective teachers to justify the four values provided.

## RESULTS

Below we analyze the prospective teachers' perceptions of sampling representativeness and variability, and justification of their responses.

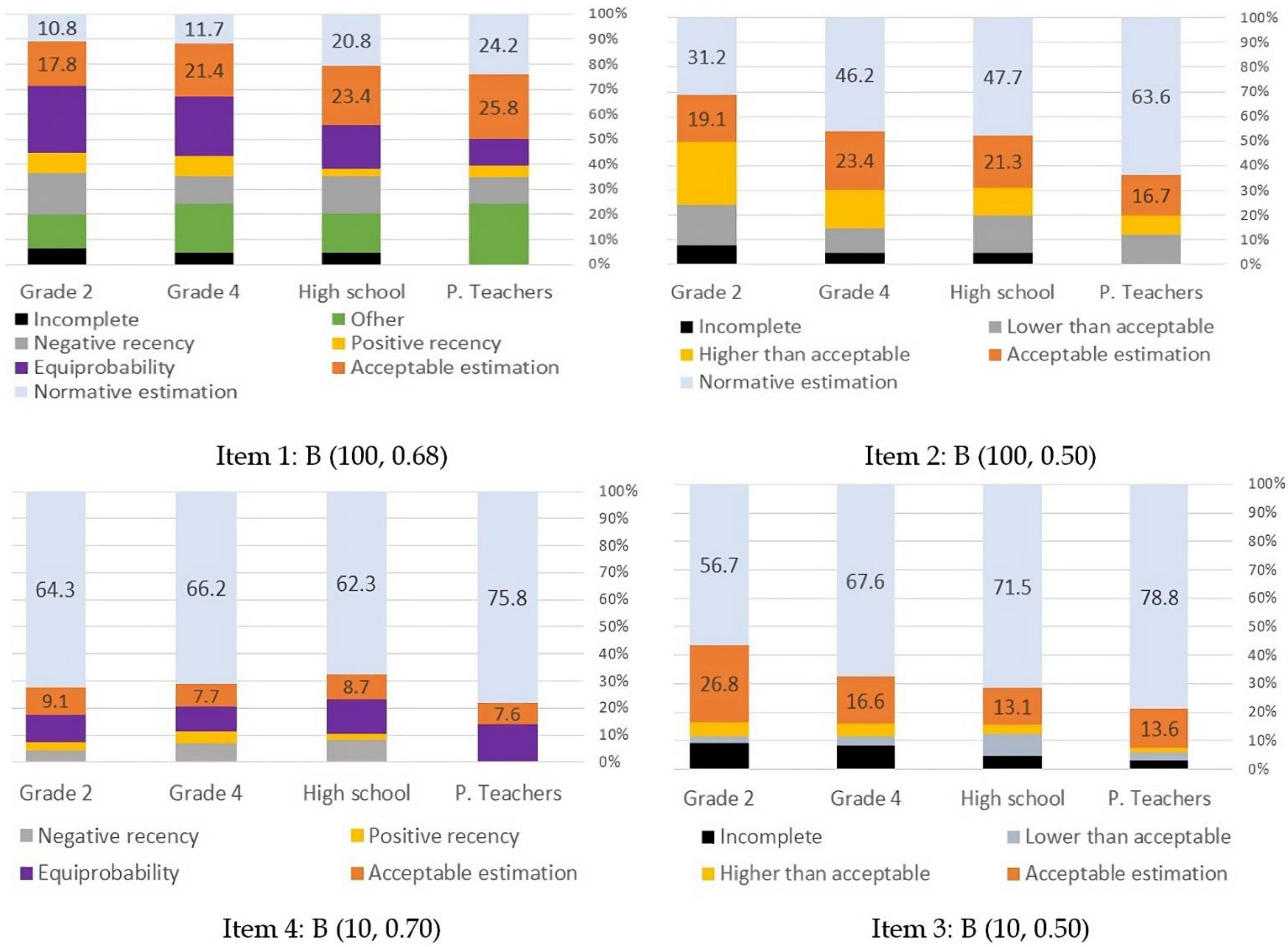


**Figure 1.** Mean of the four responses given by each participant in the items (Source: Data collected in this research)

## Perception of Sampling Representativeness

We analyzed the mean value of the four responses by the prospective teachers to each item to study the perceived sampling representativeness. These mean values were then classified into different intervals, as shown in **Figure 1**, where we have highlighted the percentages of normative (optimal) or acceptable mean values, which are the correct answers. We also compare these graphs with the results obtained in our previous research with secondary education students (Begué et al., 2018) and high school students (Batanero et al., 2020), which have been re-analyzed using the same criteria.

We observe a systematic increase in the percentage of normative and acceptable responses of students with the school year in all items and better achievement of prospective teachers. The results are much better in the tasks corresponding to the small samples, and except for the first item, the percentage of normative and acceptable responses in all groups exceeds 50%.

For the first item, a notable percentage of prospective teachers (10.6%) did not consider the experiment outcome given in the statement and provided samples with means approaching to 50%. This reasoning corresponds to the equiprobability bias (Gauvrit & Morsanyi, 2014; Lecoutre, 1992), although the item statement makes clear that the events are not equiprobable. For this item, subjects in all groups show positive recency (4.6% of prospective teachers), continuing the trend of the statement, giving much higher values than expected. Other subjects provide values that conform to negative recency (10.6% of prospective teachers). Both biases are explained by the representativeness heuristic (Tversky & Kahneman, 1974).

In item 4, 13.6% of prospective teachers provide samples whose mean value lies in the interval corresponding to equiprobability (Gauvrit & Morsanyi, 2014). Biases associated with the representativeness heuristic are also present for this item in all groups except for prospective teachers.

## Perception of Sampling Variability

We computed the range of the four values provided in each item to analyze the prospective teachers' understanding of sampling variability. We classified these ranges by considering the intervals in which the normative and acceptable estimate of the range fall (**Table 1**).

We allocated those estimates falling below the acceptable range as high concentration samples because

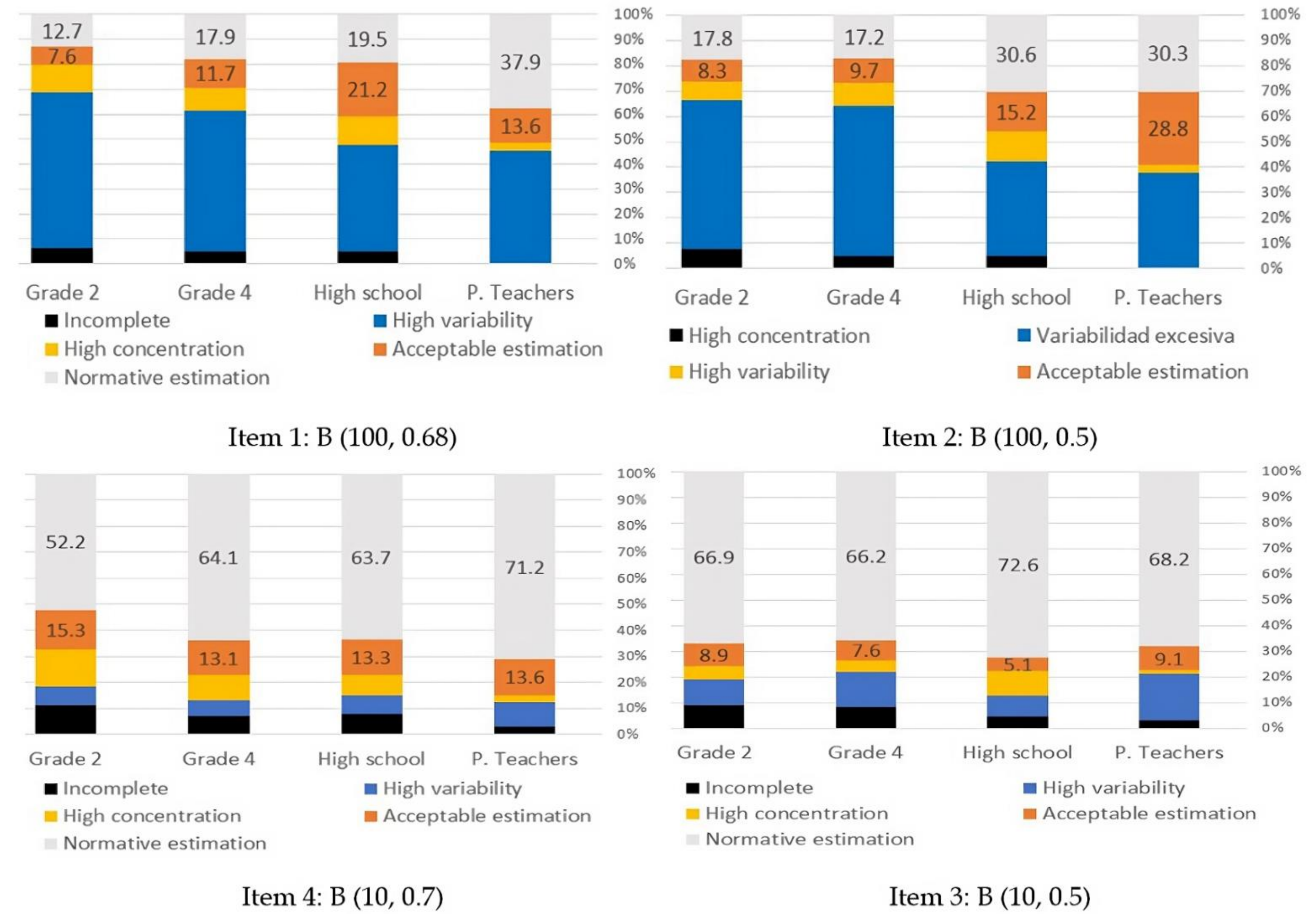


**Figure 2.** Range of the four responses given by each participant in the items (Source: Data collected in this research)

all four samples had too similar values. Similarly, we considered estimates whose range was above the acceptable interval to have high variability.

In **Figure 2**, we present the results of the prospective teachers in each item to which we add those obtained from secondary education (Begué et al., 2018) and high school (Batanero et al., 2020) students for comparison. In all items and groups, we highlight the percentages of normative and acceptable answers that are correct.

First, we observe that subjects have different perceptions of sampling variability depending on the sample size for each item, with worse results in larger samples. In items 1 and 2, a large number of participants in all groups provide high variability values relative to what is expected from the sampling distribution of the proportion. In prospective teachers, the percentage of responses with excessive variability is 45% and 38%, respectively. Thus, we infer that prospective teachers do not reach the distributional reasoning level in sampling (Shaughnessy et al., 2004), when working with large samples.

Nevertheless, the prospective teachers show a better understanding of sampling variability in the other two items compared to all groups of students. The responses improve systematically with increasing educational attainment. Consequently, there is a better understanding of sampling variability when the samples are small since the majority of participants in all groups give four values whose range is within the normative or acceptable estimation intervals.

## Justification of Responses

Batanero et al. (2020) asked 127 high school students to justify their choice of the four values in each item. The authors classified their answers according to those concepts or reasoning biases implicit in them. In the present study, we also asked the prospective teachers to explain their answers and classified their justifications according to the scheme used in Batanero et al. (2020). We describe these arguments below with transcriptions of participants' responses as examples. We denote each participant as $P_n$, where $P$ means participant and $n$ is the number assigned to his questionnaire.

1. *Randomness:* The students justify the values given based on the experiment randomness by indicating that they cannot predict the results and any outcome can occur. Furthermore, these students produce values that do not fit the item data. We find this kind of response in previous research on understanding randomness such as Briand (2005) or Savard (2010). In the following example, the prospective teacher gives that

justification for item 2, which is consistent with the values provided.

> P54, item 2: "It is entirely random. Any result is possible. The teacher's result is also random and does not condition the remaining results" (43, 65, 50, and 70).

2. *Physical properties:* The students base on the device physical characteristics, such as the drawing pin head weight (item 1) or the player's accuracy in shooting at the basket (item 4). Underlying this type of justification is the conception of probability described by Peirce (1932) as the propensity of the device or experiment to produce a specific outcome (see also Berkovitz, 2015 and Suárez, 2020).

> P21, item 1: "Because the point is less heavy than the head; therefore, although some will fall downwards, more will fall upwards" (60, 70, 63, and 74).

3. *Frequentist view of probability:* The prospective teachers estimate the probability of the event using the frequency data presented in the item statement. Thus, their responses show an understanding and correct application of the frequentist meaning of probability. In the following argument, the prospective teacher even alludes to the law of large numbers.

> P36, item 1: "As we repeat the experiment, the probability based on frequencies is similar, so if the teacher's values were 0.68 and 0.32, I assume that the students' probability would be very similar (law of large numbers)" (65, 60, 63, and 64).

4. *Classical probability:* The participant applies Laplace rule to compute the event probability as a ratio of favorable and possible cases. This justification appears especially in items 2 and 3, which describe experiments with equiprobable events, as is shown in the following example:

> P7, item 2: "Here, the results are much closer to 50% since we know that the probability of a random coin toss is 50%, i.e., by Laplace rule, we obtain that P(heads)=½=P(tail)" (50, 37, 64, and 51).

5. *Convergence and variability:* When the prospective teachers claim that the result should be similar to that given in the statement (suggesting that the idea of convergence is understood). At the same time, they indicate that there should be some variability, due to the sampling process. This type of response also appears in Gómez-Torres et al. (2016).

> P34, item 1: "The teacher got 68% of pins landing up and 32% landing down; so we expect similar percentages by the children. When moving a bit these percentages, we obtain 60-74%, but we always need to remember the weight of the pins. Therefore, it is more likely that the pin will land up" (70, 60, 72, and 65).

6. *Reasoning biases:* The response demonstrates some reasoning biases about randomness or sampling; we describe these biases in more detail in the next section.

In **Figure 3**, we present the percentages of prospective teachers and students giving each argument in their answers. Since some subjects use more than one argument, the sum of percentages is higher than 100% in each category.

First, we highlight the difference in arguments in the tasks related to equiprobable (items 2 and 3) and non-equiprobable (items 1 and 4) elementary events. In the former, most explanations refer to the classical meaning of probability, where prospective teachers apply Laplace rule, reaching over 60% of prospective teachers and over 40% of students. The trend changes in items 1 and 4, where most arguments rely on the frequentist view of probability, which is explicitly used to estimate the population proportion and to provide samples with a similar proportion. In item 4, 90% of prospective teachers and 55% of students gave this argument. We remark on a high proportion of arguments based on physical considerations in item 1. Justifications related to convergence and variability occur in noticeable proportions in all tasks, somewhat less in item 1. We note that subjects referring to these two concepts underlying the work of sampling have reached the distributional reasoning described by Shaughnessy et al. (2004).

Finally, we remark the presence of reasoning biases explicitly expressed in the arguments in all items, although infrequent, but with a higher incidence in the first task. We analyze the biases shown by the prospective teachers in the following section.

## Reasoning Biases Explicit in the Prospective Teachers' Reasoning

We have noted some reasoning biases in the justifications for the four values provided by prospective teachers. We discuss them below, with examples of the responses in which we have observed them.

1. *Equiprobability bias:* As described by Lecoutre (1992), it occurred mainly in the first task, where some participants believed that the events should be equiprobable simply because we were dealing with a random experiment. We show an example below, where, consistently with his belief in equiprobability, prospective teacher suggested values near 50% in the task.

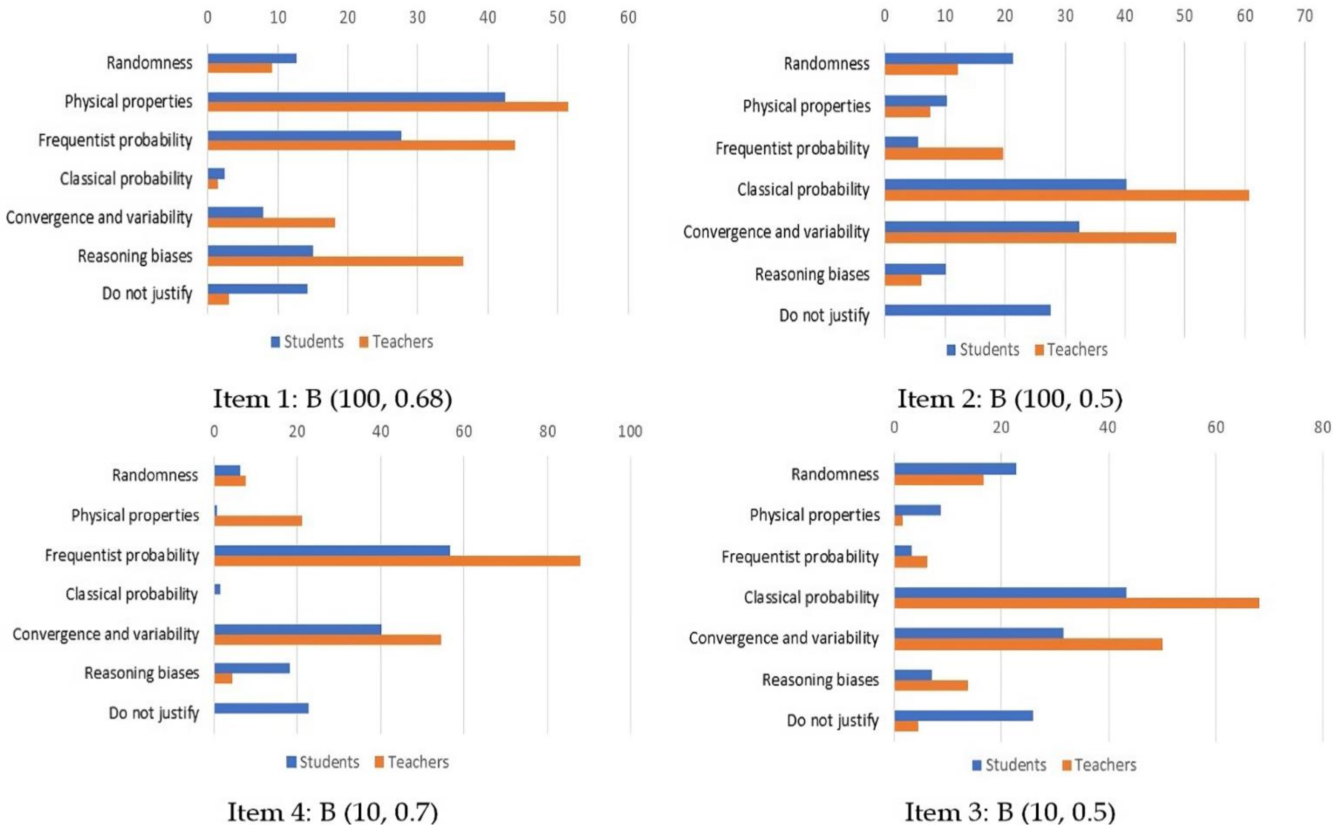


**Figure 3.** Percentage of arguments in each task in students (n=127) and prospective teachers (n=66) (Source: Data collected in this research)

> P37, item 1: "From the initial result, it seems that a drawing pin is more likely to fall upwards, but this does not mean anything, as it can be a matter of chance, which is why the results I have given tend to be 50%-50%" (40, 46, 36, and 55).

2. *Not considering independent the experiment trials:* Instead, they judged it to be a single random experiment. Thus, P54 considers throwing 100 pins as a single experiment without considering that an experiment consisting of 100 trials (observing how one pin falls each time) is mathematically equivalent to repeating the experiment consisting of dropping a pin 100 times. In both cases, the result for each drawing pin is independent of the others.

> P54, item 1: "I have 100 pins, each has a tip and a base. The probability that it will fall on one side or the other depends on its shape, and it is more likely to fall on the top, but only one trial is unreliable. So, I put a random number equal to or less than 100 since that is the number of pins" (70, 50, 100, and 60).

3. *Assuming that the situation is deterministic:* These participants interpret the item situation as deterministic without perceiving randomness. In some cases, such as the following example, the response is inconsistent by alluding to statistics and unpredictability.

> P62, item 4: "This is a deterministic (not random) phenomenon. I have given this answer to indicate that the number of baskets will remain at a high percentage, as the statistic tells us, but there is nothing that guarantees these percentages" (10, 9, 5, and 6).

4. *Outcome approach:* In this bias, described by Konold (1989), a probability question (in our case, giving four probable outcomes) is interpreted in a non-probabilistic way (indicating what will happen). Although indeed we cannot predict each experiment's outcome, there exist laws that determine the most and least probable outcomes. In the case of item 1, these would be contiguous to those obtained by the teacher.

> P1, item 1: "Any case is possible, since there is no explanation of how they stick, so they can all fall face up or face down since this is within the sample space" (50, 20, 100, and 77).

5. *Control illusion:* In this bias (Langer, 1982; Yarritu et al., 2014), subjects do not distinguish between games of skill and games of chance. They believe they can control chance and therefore have an overly high expectation of the personal

**Table 2.** Prospective teachers with reasoning biases

| Type of bias | Item 1 | Item 2 | Item 3 | Item 4 |
|---|---|---|---|---|
| 1. Equiprobability | P3, P20, P25, P37, & P47 | | P62 | |
| 2. Independence | P7, P34, P37, P56, P60, & P63 | | | P16 |
| 3. Determinism | P11 & P62 | | | P38 & P62 |
| 4. Outcome approach | P1 | P43 & P57 | P10 & P28 | |
| 5. Control illusion | P29 | | P29 | |
| 6. Representativeness | P31 | P29 | | |
| Other biases | P17 & P27 | P11 | P11, P13, P57, & P65 | |
| Total | 18 | 4 | 8 | 3 |

probability of success in the situation. This bias implies a lack of understanding of randomness, as one of the features of randomness is the impossibility of control, and it is frequently shown in compulsive gamblers. Subject P29 expresses that students can control the outcomes when describing the situation.

P29, item 1: "In this scenario, the pupils will try to overcome the pins that fell on the top but because of nerves they will get less. The last one I've put in is better than the previous as it's easier because of the calmness" (54, 32, 43, and 74).

6. *Representativeness heuristic:* When subjects suggest that the sample of four values should bear a resemblance to the population, following the reasoning described by Kahneman and Tversky (1972). In the example below, P31 gives three values similar to that of the statement to make the sample more similar to the data. Then, he provides another value to compensate for the other three results. The subject does not notice the contradiction between assuming 50% for each position and providing most values higher than that probability.

P31, item 1: "Because of the drawing pin position, I believe it is "more likely" that they will fall with the tip upwards because it is a more "stable" position. That is why there are three children where most of their pins pointing upwards and one student with reversed results, and more pins pointing downwards, although the numbers are closer to 50% of each position" (72, 61, 42, and 88).

Other reasoning biases exhibited by isolated participants included expecting a pattern in the random experiment, considering that a sample of 100 items is too small to estimate the probability, or indicating that the same results could never recur when repeating the experiment.

**Table 2** presents the participants who showed each of these biases in their responses to the different items.

Even though the number of prospective teachers who explicitly show probabilistic reasoning biases in their arguments is small, it is important to highlight these biases because they might unconsciously transmit them to their students.

## DISCUSSION

As discussed in the introduction of the paper, students need understanding of the elementary properties of sampling, due to the pervasiveness of this idea in the media and in the professional work (Burrill & Biehler, 2011). However, teaching any topic can only be successful if the teachers acquire enough mathematical teaching knowledge for that topic.

Our study adds to the literature new results related to prospective teachers' knowledge of sampling variability and representativeness, an area in which previous research is scarce. At the same time, we provide new types of tasks and methods of analysis, as regards those employed in previous research.

The findings of our study showed that prospective teachers' CCK in sampling was generally good, in coincidence with Valenzuela-Ruiz et al. (2023), and better than the students' knowledge in our previous assessment work (Begué et al., 2018; Batanero et al., 2020). The prospective teachers' perception of the sampling representativeness was good, except for the first item, where a less familiar random experiment is described. In this item, a significant proportion proposed values whose mean was almost 50% without using the frequency information in the statement. Also, some prospective teachers showed positive and negative recency (Erbas, & Ocal, 2022; Kahneman & Tversky, 1972), not perceiving the independence of the outcomes in repeated samples.

We also observed biases in the perception of too high variability in the first two items, where the sample size was high. These participants did not appreciate the relationship between the sample size and the sampling variability. Consequently, we deduce that some prospective teachers did not reach the higher (distributional) level of sampling described by Shaughnessy et al. (2004). These prospective teachers remained at the proportional level where they conceived the sample as a part of the population and related the sample proportions or means to those in the population; however, they did not consider the variability of the sampling distribution.

Moreover, although their justifications were quite complete and considered the characteristics of the proposed items, a few participants also showed sampling reasoning biases, such as equiprobability bias (Gauvrit & Morsanyi, 2014; Lecoutre, 1992), outcome approach (Konold, 1992) illusion of control (Langer, 1982; Yarritu et al., 2014) and representativeness (Erbas, & Ocal, 2022; Kahneman & Tversky, 1972).

## IMPLICATIONS FOR TEACHERS'EDUCATION

These results suggest the need to reinforce the education of prospective secondary and high school mathematics teachers on sampling, to prevent them from transmitting to their students the biases observed in their reasoning. Moreover, few prospective teachers are conscious of the students' errors in their learning of statistical inference (López-Martín et al., 2019). Thus, succeeding in the probabilistic reasoning task by teachers will improve the quality of mathematical instruction, which is usually evaluated by considering the students' outcomes (Llinares, 2021).

Another reason to enhance the education of teachers is that teachers' mathematical self-efficacy depends on their mathematical knowledge (Alshehri & Youssef, 2022). It is clear that formal mathematical training alone does not contribute to the elimination of these biases. Moreover, in addition to mathematical content knowledge the teachers need education in the different components of MKT (Ball et al., 2008; Hill et al., 2008; Phelps et al., 2020). Consequently, considering the knowledge of mathematics teachers as a style of knowing, Scheiner et al. (2019) recommend a holistic approach, instead of the separate acquirement of the components of teachers' knowledge.

Thus, to provide the participants with part of this knowledge, we recommend the activities described in this paper in workshops directed to teachers. Kula and Kocer (2020) suggest that many errors in inference learning occur because students are taught how to apply statistics and not build themselves the statistical ideas. Statistical concepts have their origin in the idea of population, adding then the idea of a sample, and finally introducing the sampling distribution. However, in applications of statistics we usually start from the sample, continue with the sampling population and finish determining the features of the population. This contradiction between construction and application of inference makes the students commit so many errors. The tasks proposed in this paper respect the logical order in the construction of sampling distribution, and therefore, follow the recommendations by Kula and Kocer (2020) to help students (in this case, prospective teachers) overcome their reasoning biases.

To make the activities more productive, after the teachers complete the tasks, the results should be discussed with them to reveal their eventual reasoning biases. In this discussion, it is necessary to emphasize the relevance of overcoming this false reasoning. Subsequently, the teacher educator should carry out simulation activities of situations raised in items using didactic software or simulators available on the Internet (e.g., binomial distribution simulation, http://www.distributome.org/js/sim/BinomialSimulation.html; simulation in a ratio: http://www.rossmanchance.com/applets/OneProp/OneProp.htm). As Koparan (2022) suggested, these activities also provide a model for teachers who should adopt simulations in their probability teaching. Moreover, according to this author, simulation activities improve the mathematical knowledge and attitudes of the teachers about probability.

In summary, the activities analyzed in the paper can contributed to simultaneously develop the prospective teachers' CCK, HCK, and some components of their didactic knowledge, in particular those referring to errors and biases in reasoning about sampling (KCS) and didactic resources for overcoming them (KCT).

**Author contributions:** All authors have sufficiently contributed to the study and agreed with the results and conclusions.

**Funding:** This study was supported under Grant PID2019-105601GB-I00 funded by MCIN/AEI/10.13039/501100011033.

**Ethical statement:** Authors stated that ethics committee approval was not required for the study. The task was part of the teaching received by the students and authorised by the centres directors and teachers. Informed consents were obtained from the participants. Data was treated as confidential information used exclusively for research purposes.

**Declaration of interest:** No conflict of interest is declared by authors.

**Data sharing statement:** Data supporting the findings and conclusions are available upon request from the corresponding author.

## REFERENCES

Alshehri, K. A., & Youssef, N. H. (2022). The influence of mathematical knowledge for teaching towards elementary teachers' mathematical self-efficacy. *EURASIA Journal of Mathematics, Science and Technology Education, 18*(6), em2118. https://doi.org/10.29333/ejmste/12086

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education, 59*(5), 389-407. https://doi.org/10.1177/0022487108324554

Batanero, C., Begué, N., Borovcnik, M., & Gea, M. (2020). Ways in which high school students understand the sampling distribution for proportions. *Statistics Education Research Journal, 19*(3), 32-52. https://doi.org/10.52041/serj.v19i3.55

Begué, N., Batanero, C., & Gea, M. M. (2018). Comprensión del valor esperado y variabilidad de la proporción muestral por estudiantes de educación secundaria obligatoria [Understanding the expected value and variability of the sample proportion by compulsory secondary education students]. *Enseñanza de las Ciencias, 36*(2), 63-79. https://doi.org/10.5565/rev/ensciencias.2256

Begué, N., Batanero, C., Ruiz, K., & Gea, M. M. (2019). Understanding sampling: A summary of the research. *Boletín de Estadística e Investigación Operativa*, *35*(1), 49-78.

Berkovitz, J. (2015). The propensity interpretation of probability: A re-evaluation. *Erkenn, 80*(Suppl 3), 629-711. https://doi.org/10.1007/s10670-014-9716-8

Briand, J. (2005). Une expérience statistique et une première approche des lois du hasard au lycée par une confrontation avec une machine simple [A statistical experiment and a first approach to the laws of chance in high school through a confrontation with a simple machine]. *Recherches en Didactique des Mathématiques* [*Research in Didactics of Mathematics*], *25*(2), 247-281.

Burrill, G. & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education* (pp. 57-69). Springer. https://doi.org/10.1007/978-94-007-1131-0

Drisko, J. W., & Maschi, T. (2016). *Content analysis.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190215491.001.0001

Eichler, A., & Vogel, M. (2014). Three approaches for modelling situations with randomness. In E. J. Chernoff, & B. Sriraman (Eds.), *Probabilistic thinking, presenting plural perspectives* (pp. 75-99). Springer. https://doi.org/10.1007/978-94-007-7155-0_4

Erbas, A. K., & Ocal, M. F. (2022). Students' intuitively-based (mis) conceptions in probability and teachers' awareness of them: the case of heuristics. *International Journal of Mathematical Education in Science and Technology*. https://doi.org/10.1080/0020739X.2022.2128454

Gauvrit N., & Morsanyi K. (2014). The equiprobability bias from a mathematical and psychological perspective. *Advances in Cognitive Psychology, 10*(4), 119-130. https://doi.org/10.5709/acp-0163-9

Gómez-Torres, E., Batanero, C., & Díaz, C. (2016). Developing a questionnaire to assess the probability content knowledge of prospective primary school teachers. *Statistics Education Research Journal, 15*(2), 197-215. https://doi.org/10.52041/serj.v15i2.248

Green, D. R. (1983). A survey of probabilistic concepts in 3000 pupils aged 11-16 years. En D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (pp. 766-783). Teaching Statistics Trust.

Harradine, A., Batanero, C., & Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education* (235–246). Springer. https://doi.org/10.1007/978-94-007-1131-0

Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics, 6*, 187-205. https://doi.org/10.1007/BF00302543

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education, 39*, 372-400. https://doi.org/10.5951/jresematheduc.39.4.0372

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgement of representativeness. *Cognitive Psychology, 3*, 430-454. https://doi.org/10.1016/0010-0285(72)90016-3

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction, 6*(1), 59-98. https://doi.org/10.1207/s1532690xci0601_3

Koparan, T. (2022). The impact of a game and simulation-based probability learning environment on the achievement and attitudes of prospective teachers. *International Journal of Mathematical Education in Science and Technology, 53*(9), 2319-2337. https://doi.org/10.1080/0020739X.2020.1868592

Kula, F., & Kocer, R. G. (2020). Why is it difficult to understand statistical inference? Reflections on the opposing directions of construction and application of inference framework. *Teaching Mathematics and its Applications, 39*(4), 248-265. https://doi.org/10.1093/teamat/hrz014

Langer, E. J. (1982). The illusion of control. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristic and biases* (pp. 231-238). Cambridge University Press. https://doi.org/10.1037/0022-3514.32.2.311

Lecoutre, M. P. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics, 23*, 557-568. https://doi.org/10.1007/BF00540060

Llinares, S. (2018). Mathematics teacher's knowledge, knowledge-based reasoning, and contexts. *Journal of Mathematics Teacher Education, 21*(1), 1-3. https://doi.org/10.1007/s10857-018-9399-1

Llinares, S. (2021). Instructional quality of mathematics teaching and mathematics teacher education. *Journal of Mathematics Teacher Education, 24*, 1-3. https://doi.org/10.1007/s10857-021-09488-2

López-Martín, M. M, Batanero, C., Díaz-Batanero, C., & Gea, M. M. (2016). La inferencia estadística en las pruebas de acceso a la universidad en Andalucía [Statistical inference in university entrance exams in Andalusia]. *Revista Paranaense de Educação Matemática, 5*(8), 33-59. https://doi.org/10.33871/22385800.2016.5.8.33-59

López-Martín, M. D. M., Batanero, C., & Gea, M. M. (2019). ¿Conocen los futuros profesores los errores de sus estudiantes en la inferencia estadística? [Do future teachers know the errors of their students in statistical inference?]. *Bolema: Boletim de Educação Matemática, 33,* 672-693. https://doi.org/10.1590/1980-4415v33n64a11

Lugo-Armenta, J. G., & Pino-Fan, L.R. (2021). Inferential statistical reasoning of math teachers: Experiences in virtual contexts generated by the COVID-19 pandemic. *Education Sciences, 11*, 363. https://doi.org/10.3390/educsci11070363

Makar, K. & Rubin, A. (2018). Learning about statistical inference. In D. Ben-Zvi, K. Makar, & J. B. Garfield (Eds.), *International handbook of research in statistics education* (pp. 261-294). Springer. https://doi.org/10.1007/978-3-319-66195-7_8

MEFP. (2022a). *Real Decreto 157/2022, de 1 de Marzo, por el que se establecen la ordenación y las enseñanzas mínimas de la educación primaria* [*Royal Decree 157/2022, of March 1, which establishes the organization and minimum teaching of primary education*]. Ministerio de Educación y Formación Profesional [Ministry of Education and Vocational Training].

MEFP. (2022b). *Real Decreto 217/2022, de 29 de Marzo, por el que se establece la ordenación y las enseñanzas mínimas de la educación secundaria obligatoria* [Royal Decree 217/2022, of March 29, which establishes the organization and minimum teaching of compulsory secondary education]. Ministerio de Educación y Formación Profesional [Ministry of Education and Vocational Training].

Peirce, C. S. (1932). Notes on the doctrine of chances. In C. S. Peirce (Ed.), *Collected papers* (pp. 404-414). Harvard University Press.

Phelps, G., Howell, H., & Liu, S. (2020). Exploring differences in mathematical knowledge for teaching for prospective and practicing teachers. *ZDM-Mathematics Education, 52*(2), 255-268. https://doi.org/10.1007/s11858-019-01097-x

Ping, C., Schellings, G., & Beijaard, D. (2018). Teacher educators' professional learning: A literature review. *Teaching and Teacher Education, 75,* 93-104. https://doi.org/10.1016/j.tate.2018.06.003

Saldanha, L., & Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics, 51*(3), 257-270. https://doi.org/10.1023/A:1023692604014

Savard, A. (2010). Simulating the risk without gambling: Can student conceptions generate critical thinking about probability? In C. Reading (Ed.), *Proceedings of the 8th International Conference on Teaching Statistics.* International Statistical Institute.

Scheiner, T., Montes, M. A., Godino, J. D., Carrillo, J., & Pino-Fan, L. R. (2019). What makes mathematics teacher knowledge specialized? Offering alternative views. *International Journal of Science and Mathematics Education, 17*(1), 153-172. https://doi.org/10.1007/s10763-017-9859-6

Shaughnessy, J. M., Ciancetta, M., & Canada, D. (2004). Types of student reasoning on sampling tasks. In M. Johnsen-Høines, & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education* (pp. 177-184). PME.

Suárez, M. (2020). *Philosophy of probability and statistical modelling.* Cambridge University Press. https://doi.org/10.1017/9781108985826

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

Valenzuela-Ruiz, S., Batanero, C., Begué, N., & Garzón-Guerrero, J. A. (2023). Conocimiento didáctico-matemático sobre la distribución de la media muestral de profesorado de bachillerato en formación [Didactic-mathematical knowledge about the distribution of the sample mean of high school teachers in training]. *Uniciencia, 37*(1), 1-20. https://doi.org/10.15359/ru.37-1.3

Wu, C., & Thompson, M. E. (2020). *Sampling theory and practice.* Springer. https://doi.org/10.1007/978-3-030-44246-0

Valenzuela-Ruiz, S., Batanero, C., Begué, N. & Garzón-Guerrero, J. A. (2023). Conocimiento didáctico-matemático sobre la distribución de la media muestral de profesorado de bachillerato en formación. *Uniciencia, 37*(1), 1-20. https://doi.org/10.15359/ru.37-1.3

Yarritu, I., Matute, H., & Vadillo, M. A. (2014). Illusion of control: The role of personal involvement. *Experimental Psychology, 61*(1), 38-47. https://doi.org/10.1027/1618-3169/a000225

## APPENDIX: QUESTIONNAIRE

### Task 1

A parcel of 100 drawing pins is emptied out onto a table by a teacher. Some drawing pins landed "up" and some landed "down". The results were as follows: 68 landed up and 32 landed down. The teacher then asked four students to repeat the experiment (with the same pack of drawing pins). Each student emptied the pack of 100 drawing pins and got some landing up and some landing down.

In **Table A1**, write one probable result for each student.

**Table A1.**

| Daniel | Martín | Diana | María |
|---|---|---|---|
| Up: | Up: | Up: | Up: |
| Down: | Down: | Down: | Down: |

### Task 2

A parcel of 100 fair coins is emptied onto a table by a teacher with the following result: 53 Heads and 47 Tails. The teacher then asked four students to repeat the experiment with the same parcel of coins. Each student emptied the pack of 100 fair coins and got some heads and some tails. In **Table A2**, write one probable result for each student.

**Table A2.**

| Elena | Clara | Matías | Rosa |
|---|---|---|---|
| Heads: | Heads: | Heads: | Heads: |
| Tails: | Tails: | Tails: | Tails: |

### Task 3

A teacher asks 4 students to flip 10 coins on the table and count the number of Heads and Tails. In **Table A3**, write one probable result for each student.

**Table A3.**

| Silvia | Javier | Miguel | Carmen |
|---|---|---|---|
| Heads: | Heads: | Heads: | Heads: |
| Tails: | Tails: | Tails: | Tails: |

### Task 4

From 100 attempts of a basket-ball player to throw the ball into the basket from the free-throw line, 70 are shots (land in the basket). In **Table A4**, write one probable result for four games in which he throws the ball 10 times from the free-throw line.

**Table A4.**

| Game 1 (10 throws) | Game 2 (10 throws) | Game 3 (10 throws) | Game 4 (10 throws) |
|---|---|---|---|
| Number of shots: | Number of shots: | Number of shots: | Number of shots: |
| Number of failures: | Number of failures: | Number of failures: | Number of failures: |